Arch Appl Mech (2006) 76: 327–348
DOI 10.1007/s00419-006-0031-y

ORIGINAL

C. Studer · Ch. Glocker

Representation of normal cone inclusion problems in dynamics via non-linear equations

Received: 8 november 2005 / Accepted: 5 april 2006 / Published online: 12 August 2006

Abstract Analyzing non-smooth mechanical systems requires often the solution of inclusion problems of normal cone type. These problems arise for example in the event-driven or time-stepping simulation approaches. Such inclusion problems can be written as non-linear equations, which can be solved iteratively. In this paper we discuss three different methods to derive the non-linear equations representing the inclusions arising in the event-driven simulation approach. First, we formulate inclusions describing the individual non-smooth constraints and solve them successively. Secondly, we interpret the non-linear equations as the conditions for the saddle point of the augmented Lagrangian function. As a third possibility we discuss the exact regularization of set-valued force laws. All three methods lead to the same numerical scheme, but give different insight into the problem. Especially the factor r occurring in the non-linear equations is discussed. Two iterative methods for solving the non-linear equations are presented together with some remarks on convergence.

Keywords non-smooth dynamics · augmented Lagrangian · exact regularization · contact problems · inclusions

1 Introduction

A non-smooth mechanical system with unilateral contacts and friction can be treated either by event-driven methods [6] or time-stepping methods [8,11,13]. In both cases, inclusions have to be solved: in the case of the event-driven methods to determine the future contact behaviour and to evaluate impacts, in the case of the time-stepping algorithms to solve the time-stepping inclusions. We refer in this paper to the event-driven inclusions, which describe the contact behaviour on acceleration level. These inclusions are often solved by setting up linear or non-linear complementarity problems [6,8], or by turning them into non-linear equations which can be solved by iterative methods [1,9,13]. We discuss three different methods to obtain these non-linear equations. First, we merge the equations of motion and the set-valued force laws to obtain inclusions that represent the different non-smooth constraints. Based on the solution of a single inclusion which describes one individual non-smooth constraint, a successive solution technique can be found, which corresponds to a Jacobi or Gauss–Seidel like iterative method. Secondly, we derive the non-linear equations from the conditions for the saddle point of the augmented Lagrangian function, which replaces a constrained optimization problem by an unconstrained [2,9,10]. A third possibility to derive these non-linear equations is exact regularization [9], for which the set-valued force laws are regularized in a manner that both, the solution of the original and the regularized problem coincide with each other.

C. Studer (✉) · Ch. Glocker
IMES-Center of Mechanics,
ETH Zurich,
8092 Zurich, Switzerland
E-mail: christian.studer@imes.mavt.ethz.ch,
christoph.glocker@imes.mavt.ethz.ch

The aim of this paper is to widen the reader's view on the subject by discussing how the different solution concepts are interconnected. In particular, various interpretations of the parameter r occurring in the non-linear equations are discussed.

2 The contact problem on acceleration level

The purpose of this paper is to show how inclusions originating from set-valued force laws can be written as non-linear equations for finite-dimensional second order dynamical systems. To demonstrate the different methods, we have exemplarily chosen the dynamical take-off and stick-slip transition of frictional contacts within the event-driven approach. Impacts as well as contacts in pure sliding regime are intentionally left out for conciseness, but can be treated in a similar way. From the viewpoint of dynamics, we aim at determining the right accelerations of the system, i.e. the right derivatives $\dot{\mathbf{u}}(t)$ of the generalized velocities $\mathbf{u}(t)$ for a given post-impact state $(\mathbf{q}(t), \mathbf{u}(t))$ at a fixed time t. The positions of the system at time t are addressed by the generalized coordinates $\mathbf{q}(t)$ with $\dot{\mathbf{q}}(t) = \mathbf{u}(t)$ almost everywhere. Our approach to impose frictional contact constraints on the system is lead by the following philosophy [6–8]: Contact forces $\boldsymbol{\lambda}$ are taken into account in the equations of motion by Lagrangian multipliers. Contact laws are expressed as set-valued constitutive laws in local contact coordinates on acceleration level, $-\boldsymbol{\lambda} \in D(\dot{\gamma})$. Kinematic transformations $\dot{\mathbf{u}} \to \dot{\gamma}$ are used to finally state the whole problem in contact entities, $\dot{\gamma} = \mathbf{G}\boldsymbol{\lambda} + \mathbf{c}$, and to solve this equation together with the set-valued force laws.

2.1 Set-valued force laws

Non-smooth constraints can be described by set-valued force laws [7]. They are often expressed in the framework of convex analysis, which gives a very compact view on the subject. In this paper we try to avoid using convex analysis to make it accessible also for readers not so familiar with this subject. We call a constraint which prevents bodies to penetrate each other a unilateral constraint. A force law modeling frictional behaviour is called a frictional constraint. We denote the i-th constraint by the index i which can either be a unilateral constraint or a frictional constraint. A unilateral constraint is characterized by the normal contact force λ_i and the (normal) gap function $g_i(\mathbf{q})$ with admissible values $g_i \geq 0$. The relative velocity in normal direction is denoted by $\gamma_i(t) = \dot{g}_i(t)$, and the associated accelerations by $\dot{\gamma}_i(t)$ with

$$\dot{\gamma}_i = \mathbf{w}_i^\top \dot{\mathbf{u}} + \zeta_i. \tag{1}$$

It can be shown that $\mathbf{w}_i$ coincides with the generalized force direction of λ_i, if g_i measures the Euclidean contact distance and λ_i acts in the same direction as the one specified by g_i. To model the take-off transition on acceleration level, we introduce the set of unilateral contacts which are active on velocity level at time t, $P_N(t) = \{i \mid g_i(\mathbf{q}) = 0, \gamma_i(\mathbf{q}, \mathbf{u}) = 0\}$, and impose for every $i \in P_N$ a constitutive law of the form

$$-\lambda_i \in \mathrm{Upr}(\dot{\gamma}_i) \Leftrightarrow \begin{cases} \dot{\gamma}_i > 0 \Rightarrow \lambda_i = 0 \\ \dot{\gamma}_i = 0 \Rightarrow \lambda_i \geq 0 \end{cases} \tag{2}$$

with inverse

$$-\dot{\gamma}_i \in \mathrm{Upr}^{-1}(\lambda_i) = \mathrm{Upr}(\lambda_i) \Leftrightarrow \begin{cases} \lambda_i > 0 \Rightarrow \dot{\gamma}_i = 0 \\ \lambda_i = 0 \Rightarrow \dot{\gamma}_i \geq 0\,. \end{cases} \tag{3}$$

Spatial friction requires the definition of a tangent contact plane, which shall be spanned by two orthonormalized vectors $(\mathbf{t}_1, \mathbf{t}_2)$. The coordinates of the tangential contact force and the tangential contact relative velocity with respect to this basis are denoted by $\boldsymbol{\lambda}_i$ and $\gamma_i(\mathbf{q}, \mathbf{u})$, respectively. The tangential contact relative acceleration $\dot{\gamma}_i$ may be expressed in the form

$$\dot{\gamma}_i = \mathbf{W}_i^\top \dot{\mathbf{u}} + \boldsymbol{\zeta}_i, \tag{4}$$

in which the two columns of $\mathbf{W}_i$ correspond to the two generalized force directions associated with $\boldsymbol{\lambda}_i$. To model the stick-slip transition on acceleration level, we introduce the set $P_{T3} = \{i | \gamma_i(\mathbf{q}, \mathbf{u}) = 0\}$ of active stick constraints on velocity level and impose, for every $i \in P_{T3}$, the spatial law of dry friction

$$-\boldsymbol{\lambda}_i \in a_i \mathrm{SgnSp}(\dot{\gamma}_i) \Leftrightarrow \begin{cases} ||\dot{\gamma}_i|| > 0 \Rightarrow \boldsymbol{\lambda}_i = -a_i \frac{\dot{\gamma}_i}{||\dot{\gamma}_i||} \\ \dot{\gamma}_i = 0 \quad \Rightarrow ||\boldsymbol{\lambda}_i|| \leq a_i. \end{cases} \tag{5}$$

Note that a_i is the maximum friction force of the i-th friction constraint. If the friction constraint i belongs to the unilateral contact $j \in P_N$, then $a_i = \mu\lambda_j$. The inverse set-valued force law for a spatial friction constraint i is

$$-\dot{\gamma}_i \in \mathrm{SgnSp}^{-1}\left(\frac{\boldsymbol{\lambda}_i}{a_i}\right) \Leftrightarrow \begin{cases} ||\boldsymbol{\lambda}_i|| = a_i \Rightarrow \dot{\gamma}_i \in -\mathbb{R}_0^+ \ \frac{\boldsymbol{\lambda}_i}{||\boldsymbol{\lambda}_i||} \\ ||\boldsymbol{\lambda}_i|| < a_i \Rightarrow \dot{\gamma}_i = 0. \end{cases} \tag{6}$$

2.2 Equations of motion

We first consider a finite-dimensional second order dynamical system for the case that no contact forces are present. We denote by $\mathbf{q}$ the generalized coordinates of this system and by $\mathbf{u} = \dot{\mathbf{q}}$ the associated generalized velocities. Lagrange's equations of second kind yield the equations of motion, which we write in the form $\mathbf{M}\dot{\mathbf{u}} - \mathbf{h} = 0$ with $\mathbf{M} = \mathbf{M}(\mathbf{q})$ being the symmetric and positive definite mass matrix and $\mathbf{h} = \mathbf{h}(\mathbf{q}, \mathbf{u})$ the vector of all external and gyroscopic forces acting on the system. Additional contact forces are taken into account by Lagrangian multipliers,

$$\mathbf{M}\dot{\mathbf{u}} - \mathbf{h} - \sum_{i=1}^{n} \mathbf{W}_i \boldsymbol{\lambda}_i = 0. \tag{7}$$

by which the set-valued constitutive laws from the last section are now connected to the system.

2.3 Contact inclusions on acceleration level

In this section we transform the entire problem to contact coordinates and state it as an inclusion likewise in terms of the unknown contact forces $\boldsymbol{\lambda}$ or contact accelerations $\dot{\gamma}$. By defining global vectors and matrices

$$\dot{\gamma} = \begin{pmatrix} \dot{\gamma}_1 \\ \vdots \\ \dot{\gamma}_n \end{pmatrix}, \quad \boldsymbol{\lambda} = \begin{pmatrix} \boldsymbol{\lambda}_1 \\ \vdots \\ \boldsymbol{\lambda}_n \end{pmatrix}, \quad \mathbf{W} = \begin{pmatrix} \mathbf{W}_1 \cdots \mathbf{W}_n \end{pmatrix}, \quad \boldsymbol{\zeta} = \begin{pmatrix} \boldsymbol{\zeta}_1 \\ \vdots \\ \boldsymbol{\zeta}_n \end{pmatrix}, \tag{8}$$

the equations of motion (7) are solved for $\dot{\gamma}$ and put into the kinematic transformations (1) and (4), which yields

$$\dot{\gamma} = \mathbf{W}^\top \mathbf{M}^{-1} \mathbf{W} \boldsymbol{\lambda} + \mathbf{W}^\top \mathbf{M}^{-1} \mathbf{h} + \boldsymbol{\zeta} =: \mathbf{G}\boldsymbol{\lambda} + \mathbf{c}. \tag{9}$$

Note that the matrix $\mathbf{G}$ is positive definite if the force directions $\mathbf{W}_i$ are independent, otherwise positive semidefinite. Due to the positive definite mass matrix $\mathbf{M}$, all diagonal entries G_{ii} of $\mathbf{G}$ are greater than zero. Equation (9) forms a linear system, in which the rows represent the non-smooth constraints; a unilateral contact is represented by one row, a spatial friction constraint by two rows. Together with the associated set-valued force laws, we arrive at n inclusions describing the n non-smooth constraints. If the constraint i is a unilateral contact ($i \in P_N$), then the inclusion resulting from (2) and (9) is

$$\sum_{j=1}^{n} (\mathbf{G}^{-1})_{ij} \dot{\gamma}_j + \mathrm{Upr}(\dot{\gamma}_i) \ni (\mathbf{G}^{-1}\mathbf{c})_i, \tag{10}$$

or, by using the inverse representation (3) of (2),

$$\sum_{j=1}^{n} \mathbf{G}_{ij} \boldsymbol{\lambda}_j + \mathrm{Upr}(\lambda_i) \ni -c_i. \tag{11}$$

If the constraint i is a spatial friction constraint ($i \in P_{T3}$), then the resulting inclusion is by (5) and (9)

$$\sum_{j=1}^{n} (\mathbf{G}^{-1})_{ij} \dot{\gamma}_j + a_i \mathrm{SgnSp}(\dot{\gamma}_i) \ni (\mathbf{G}^{-1}\mathbf{c})_i, \tag{12}$$

or, by using the inverse (6),

$$\sum_{j=1}^{n} \mathbf{G}_{ij}\boldsymbol{\lambda}_j + \mathrm{SgnSp}^{-1}\left(\frac{\boldsymbol{\lambda}_i}{a_i}\right) \ni -\mathbf{c}_i. \tag{13}$$

By setting up these inclusions for all non-smooth constraints, the contact behaviour of the non-smooth mechanical system is completely determined. The inclusions (10) and (12) describe the contact behaviour in terms of the relative contact accelerations $\dot{\gamma}_i$. We call them $\dot{\gamma}$-inclusions. The inclusions (11) and (13) describe the contact behaviour in terms of the constraint forces $\boldsymbol{\lambda}_i$ and are therefore called $\boldsymbol{\lambda}$-inclusions.

2.4 Example

In this example we show how a system with one unilateral and one spatial friction constraint has to be treated. For this case, Eq. (9) together with the contact laws (3) and (6) reads

$$\begin{pmatrix}\dot{\gamma}_1\\ \dot{\gamma}_2\end{pmatrix} = \begin{pmatrix}G_{11} & \mathbf{G}_{12}\\ \mathbf{G}_{21} & \mathbf{G}_{22}\end{pmatrix}\begin{pmatrix}\lambda_1\\ \boldsymbol{\lambda}_2\end{pmatrix} + \begin{pmatrix}c_1\\ \mathbf{c}_2\end{pmatrix} \tag{14}$$

$$-\dot{\gamma}_1 \in \mathrm{Upr}(\lambda_1) \tag{15}$$

$$-\dot{\gamma}_2 \in \mathrm{SgnSp}^{-1}\left(\frac{\boldsymbol{\lambda}_2}{a_2}\right) \tag{16}$$

Please note that both, the relative contact acceleration $\dot{\gamma}_1$ and the contact force λ_1 of the unilateral contact are one-dimensional scalars. On the other hand, the relative contact acceleration $\dot{\gamma}_2$ and the contact force $\boldsymbol{\lambda}_2$ of the spatial friction constraint are two-dimensional vectors. The vector $\dot{\gamma} = (\dot{\gamma}_1\ \dot{\gamma}_2)^\top$ of all relative contact accelerations and the vector $\boldsymbol{\lambda} = (\lambda_1\ \boldsymbol{\lambda}_2)^\top$ of all contact forces are thus three-dimensional vectors. The matrix $\mathbf{G}$ has dimensions 3×3, and the partial matrices $\mathbf{G}_{ij}$ are

$$G_{11} \in \mathbb{R}^{1\times 1}, \quad \mathbf{G}_{12} \in \mathbb{R}^{1\times 2}, \quad \mathbf{G}_{21} \in \mathbb{R}^{2\times 1}, \quad \mathbf{G}_{22} \in \mathbb{R}^{2\times 2}. \tag{17}$$

Elimination of $\dot{\gamma}_1$ and $\dot{\gamma}_2$ from (14), (15), (16) yields the two $\boldsymbol{\lambda}$-inclusions

$$\sum_{j=1}^{2} \mathbf{G}_{1j}\boldsymbol{\lambda}_j + \mathrm{Upr}(\lambda_1) \ni -c_1, \tag{18}$$

$$\sum_{j=1}^{2} \mathbf{G}_{2j}\boldsymbol{\lambda}_j + \mathrm{SgnSp}^{-1}\left(\frac{\boldsymbol{\lambda}_2}{a_2}\right) \ni -\mathbf{c}_2, \tag{19}$$

which completely determine the problem.

3 Analysis of one single non-smooth constraint

In this section we introduce the functions prox_C and the κ_C. Both functions are continuous and monotone, and will later be used to rewrite inclusions (10), (11), (12), (13) as equivalent *equations*. The index C denotes the convex set on which the functions prox_C and κ_C act. We use the set $C = \mathbb{R}_0^+$ for unilateral contacts and the set S

$$S = \{\boldsymbol{\xi} \in \mathbb{R}^2 |\ ||\boldsymbol{\xi}||_2 \le a\} \tag{20}$$

for the spatial friction constraints. The function prox_C performs a projection onto the set C, that is the point $\mathbf{x} = \text{prox}_C(\boldsymbol{\xi})$ is the nearest point to $\boldsymbol{\xi}$ in the set C (proximal point). The functions prox_C for the selected sets C are

$$C = \mathbb{R}_0^+ \;\Rightarrow\; \text{prox}\,\mathbb{R}_0^+(\xi) = \begin{cases} \xi & \text{if } \xi \geq 0 \\ 0 & \text{if } \xi < 0 \end{cases} \tag{21}$$

$$C = S \;\Rightarrow\; \text{prox}_S(\boldsymbol{\xi}) = \begin{cases} \boldsymbol{\xi} & \text{if } \boldsymbol{\xi} \in S \\ \frac{\boldsymbol{\xi}}{||\boldsymbol{\xi}||_2}\, a & \text{if } \boldsymbol{\xi} \notin S. \end{cases} \tag{22}$$

The function κ_C for the set S is defined as

$$C = S \;\Rightarrow\; \kappa_S(\boldsymbol{\xi}, r) = \begin{cases} \boldsymbol{\xi} - r\, a \frac{\boldsymbol{\xi}}{||\boldsymbol{\xi}||_2} & \text{if } \frac{1}{r}\,\boldsymbol{\xi} \notin S \\ 0 & \text{if } \frac{1}{r}\,\boldsymbol{\xi} \in S\,. \end{cases} \tag{23}$$

With the help of the functions prox_C and κ_C we may now rewrite the inclusions (10), (11), (12), (13) needed in the contact problem as *equations*. In particular, for $r > 0$ it holds true that

$$x_i + r\,\text{Upr}(x_i) \ni b_i \quad \Leftrightarrow \quad x_i = \text{prox}_{\mathbb{R}_0^+}(b_i), \tag{24}$$

$$\mathbf{x}_i + r\, a_i\,\text{SgnSp}(\mathbf{x}_i) \ni \mathbf{b}_i \quad \Leftrightarrow \quad \mathbf{x}_i = \kappa_S(\mathbf{b}_i, r), \tag{25}$$

$$\mathbf{x}_i + r\,\text{SgnSp}^{-1}(\frac{\mathbf{x}_i}{a_i}) \ni \mathbf{b}_i \quad \Leftrightarrow \quad \mathbf{x}_i = \text{prox}_S(\mathbf{b}_i). \tag{26}$$

The associated proofs are performed within the framework of convex analysis and are presented in appendix. Most important to note is that non-smooth constraints may now be formulated as equations, which we explicitly show by the examples of a unilaterally constrained particle and a particle with friction.

3.1 Unilateral contact

An active unilateral contact between a point mass and a surface (Fig. 1) is described by the $\boldsymbol{\lambda}$-inclusion

$$\left.\begin{array}{l} m\dot{u} = \lambda - F \\ -\dot{u} \in \text{Upr}(\lambda) \end{array}\right\} \quad \Leftrightarrow \quad \lambda + m\text{Upr}(\lambda) \ni F \tag{27}$$

in which F is the external forcing and $\dot{x} = u$ the velocity of the particle. With the help of (24), the $\boldsymbol{\lambda}$-inclusion in (27) may be replaced by a prox condition and solved according to (21),

$$\lambda = \text{prox}\,\mathbb{R}_0^+(F) \Leftrightarrow \begin{cases} \lambda = F & \text{if } F \in \mathbb{R}_0^+ \\ \lambda = 0 & \text{if } F \notin \mathbb{R}_0^+. \end{cases} \tag{28}$$

Note that the set-valued force laws are formulated on acceleration level and that we have assumed an active contact on velocity level ($x = 0,\ u = 0$).

3.2 Spatial friction constraint

For an active spatial friction constraint ($u = 0$) between a point mass and a surface (Fig. 2), the $\dot{\boldsymbol{\gamma}}$-inclusion is

$$\left.\begin{array}{l} m\dot{\mathbf{u}} = \boldsymbol{\lambda} + \mathbf{F} \\ -\boldsymbol{\lambda} \in a\text{SgnSp}(\dot{\mathbf{u}}) \end{array}\right\} \quad \Leftrightarrow \quad \dot{\mathbf{u}} + \frac{1}{m}\, a\,\text{SgnSp}(\dot{\mathbf{u}}) \in \frac{\mathbf{F}}{m}\,. \tag{29}$$

By taking into account (25), we obtain

$$\dot{\mathbf{u}} = \kappa_S\left(\frac{\mathbf{F}}{m}, \frac{1}{m}\right) \Leftrightarrow \begin{cases} \dot{\mathbf{u}} = 0 & \text{if } \mathbf{F} \in S \\ \dot{\mathbf{u}} = \frac{\mathbf{F}}{m} - \frac{a}{m}\,\frac{\mathbf{F}}{||\mathbf{F}||_2} & \text{if } \mathbf{F} \notin S. \end{cases} \tag{30}$$

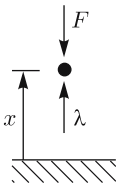
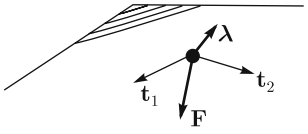


Fig. 1 Unilateral constraint: the position of a point mass is described by the coordinate x. In addition to the load F, a contact force λ is introduced



$$\begin{aligned} \mathbf{F} &= F_1\mathbf{t}_1 + F_2\mathbf{t}_2 \\ \boldsymbol{\lambda} &= \lambda_1\mathbf{t}_1 + \lambda_2\mathbf{t}_2 \\ \mathbf{q} &= q_1\mathbf{t}_1 + q_2\mathbf{t}_2 \end{aligned}$$

Fig. 2 Spatial friction constraint: point mass lying on a surface, with external force $\mathbf{F}$ and friction force $\boldsymbol{\lambda}$. The two forces are described in a tangential coordinate system ($\mathbf{t_1}$, $\mathbf{t_2}$)

The inverse problem of (29) leads to the $\boldsymbol{\lambda}$-inclusion

$$\left.\begin{array}{ll} m\dot{\mathbf{u}} & = \boldsymbol{\lambda} + \mathbf{F} \\ -\dot{\mathbf{u}} \in \mathrm{SgnSp}^{-1}\left(\frac{\boldsymbol{\lambda}}{a}\right) & \end{array}\right\} \quad \Leftrightarrow \quad \boldsymbol{\lambda} + m\,\mathrm{SgnSp}^{-1}\left(\frac{\boldsymbol{\lambda}}{a}\right) \in -\mathbf{F}\,, \tag{31}$$

which can be solved by (26) to give

$$\boldsymbol{\lambda} = \mathrm{prox}_S(-\mathbf{F}) \Leftrightarrow \begin{cases} \boldsymbol{\lambda} = -\mathbf{F} & \text{if } -\mathbf{F} \in S \\ \boldsymbol{\lambda} = -a\,\frac{\mathbf{F}}{||\mathbf{F}||_2} & \text{if } -\mathbf{F} \notin S. \end{cases} \tag{32}$$

Note that the set-valued force laws are formulated on acceleration level and that we have assumed that stick is active on velocity level ($\mathbf{u} = 0$).

4 Representing contact inclusions by non-linear equations

With the help of (24), (25), (26) the inclusions (10), (11), (12), (13) describing the contact behaviour can now be written as equations. The $\boldsymbol{\lambda}$-inclusion (11) for a unilateral constraint $i \in P_N$ becomes

$$-\sum_{\substack{j=1 \\ j\neq i}}^{n} \mathbf{G}_{ij}\,\boldsymbol{\lambda}_j - c_i \in G_{ii}\lambda_i + \mathrm{Upr}(\lambda_i). \tag{33}$$

Since $G_{ii} > 0$ is a scalar, we obtain together with (24)

$$\lambda_i = \mathrm{prox}_{\mathbb{R}_0^+}\left(-\sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{\mathbf{G}_{ij}}{G_{ii}}\,\boldsymbol{\lambda}_j - \frac{c_i}{G_{ii}}\right). \tag{34}$$

The force λ_i of the i-th non-smooth constraint appears only on the left hand side of (34). Thus it can be directly computed if all other forces $\boldsymbol{\lambda}_j$ are known. The quotient $\mathbf{G}_{ij}/G_{ii}$ shows the influence of the j-th contact on the i-th contact. A non-linear equation in which the unknowns are the accelerations $\dot{\gamma}$ follows from the $\dot{\gamma}$-inclusion (10),

$$\dot{\gamma}_i = \mathrm{prox}_{\mathbb{R}_0^+}\left(-\sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{(\mathbf{G}^{-1})_{ij}}{(\mathbf{G}^{-1})_{ii}}\,\dot{\gamma}_j + \frac{(\mathbf{G}^{-1}\mathbf{c})_i}{(\mathbf{G}^{-1})_{ii}}\right), \tag{35}$$

in which $(\mathbf{G}^{-1})_{ii} > 0$ is a scalar.

The $\boldsymbol{\lambda}$-inclusion (13) characterizing a spatial friction constraint $i \in P_{T3}$ can be reformulated as

$$-\sum_{\substack{j=1\\ j\neq 1}}^{n} \mathbf{G}_{ij}\,\boldsymbol{\lambda}_j - \mathbf{c}_i \in \mathbf{G}_{ii}\boldsymbol{\lambda}_i + \mathrm{SgnSp}^{-1}(\frac{\boldsymbol{\lambda}_i}{a_i}). \tag{36}$$

In this case, $\mathbf{G}_{ii}$ is no longer a scalar, but a 2×2 matrix. By splitting it up according to

$$\mathbf{G}_{ii} = \alpha_i \mathbf{I} + \mathbf{B}_{ii} \tag{37}$$

with $\mathbf{I}$ the identity matrix, $\alpha_i > 0$, and $\mathbf{B}_{ii}$ the remaining part, one obtains

$$-\sum_{\substack{j=1\\ j\neq 1}}^{n} \mathbf{G}_{ij}\,\boldsymbol{\lambda}_j - \mathbf{c}_i - \mathbf{B}_{ii}\boldsymbol{\lambda}_i \in \alpha_i\boldsymbol{\lambda}_i + \mathrm{SgnSp}^{-1}\left(\frac{\boldsymbol{\lambda}_i}{a_i}\right). \tag{38}$$

This yields, together with (26), the non-linear equation

$$\boldsymbol{\lambda}_i = \mathrm{prox}_S\left(-\sum_{\substack{j=1\\ j\neq i}}^{n} \frac{\mathbf{G}_{ij}}{\alpha_i}\,\boldsymbol{\lambda}_j - \frac{\mathbf{c}_i}{\alpha_i} - \frac{\mathbf{B}_{ii}}{\alpha_i}\,\boldsymbol{\lambda}_i\right). \tag{39}$$

Note that the force $\boldsymbol{\lambda}_i$ of the i-th non-smooth constraint cannot be determined straightforward as function of $\boldsymbol{\lambda}_j$ because it appears on both sides of the equation due to the remainder matrix $\mathbf{B}_{ii}$. This remainder matrix $\mathbf{B}_{ii}$ depends on the choice of the the tangential unit vectors $\mathbf{t}_1$ and $\mathbf{t}_2$ of the spatial friction constraint, which can sometimes be chosen such that $\mathbf{B}_{ii}$ becomes zero. Note that $\mathbf{B}_{ii} = 0$ is quite handy but not required to solve the problem. The non-linear equation describing the spatial friction constraint in terms of the accelerations $\dot{\gamma}$ follows from the $\dot{\gamma}$-inclusion (12),

$$\dot{\gamma}_i = \kappa_S\left(-\sum_{\substack{j=1\\ j\neq i}}^{n} \frac{\mathbf{G}^{-1}_{ij}}{\alpha_i}\,\dot{\gamma}_j + \frac{(\mathbf{G}^{-1}\mathbf{c})_i}{\alpha_i} - \frac{\mathbf{B}_{ii}}{\alpha_i}\,\dot{\gamma}_i\,,\ \frac{1}{\alpha_i}\right). \tag{40}$$

A more general formulation of the non-linear equations (34) and (39) can be found by first rearranging the $\boldsymbol{\lambda}$-inclusions (11) and (13) in the form

$$-\sum_{j=1}^{n} \mathbf{G}_{ij}\,\boldsymbol{\lambda}_j - c_i + \frac{\lambda_i}{r_i} \in \frac{\lambda_i}{r_i} + \mathrm{Upr}(\lambda_i) \qquad i \in P_N, \tag{41}$$

$$-\sum_{j=1}^{n} \mathbf{G}_{ij}\,\boldsymbol{\lambda}_j - \mathbf{c}_i + \frac{\boldsymbol{\lambda}_i}{r_i} \in \frac{\boldsymbol{\lambda}_i}{r_i} + \mathrm{SgnSp}^{-1}(\frac{\boldsymbol{\lambda}_i}{a_i}) \qquad i \in P_{T3}. \tag{42}$$

Those inclusions are then multiplied by r_i, and relations (24) and (26) are used to obtain the non-linear equations

$$\lambda_i = \mathrm{prox}_{\mathbb{R}_0^+}\left(\lambda_i - r_i\left(\sum_{j=1}^{n} \mathbf{G}_{ij}\,\boldsymbol{\lambda}_j + c_i\right)\right) \qquad i \in P_N, \tag{43}$$

$$\boldsymbol{\lambda}_i = \mathrm{prox}_S\left(\boldsymbol{\lambda}_i - r_i\left(\sum_{j=1}^{n} \mathbf{G}_{ij}\,\boldsymbol{\lambda}_j + \mathbf{c}_i\right)\right) \qquad i \in P_{T3}. \tag{44}$$

The factor r_i is an arbitrary positive scalar which can be chosen such that (41), (42) become (34), (39). An interpretation of r_i is given in the following sections. The same can be carried out for the $\dot{\gamma}$ equations (35) and (40), which yields the more general formulation

$$\dot{\gamma}_i = \operatorname{prox}_{\mathbb{R}_0^+}\left(\gamma_i - r_i\left(\sum_{j=1}^{n}\left(\mathbf{G}^{-1}\right)_{ij}\dot{\gamma}_j - \left(\mathbf{G}^{-1}\mathbf{c}\right)_i\right)\right) \qquad i \in P_N, \tag{45}$$

$$\dot{\gamma}_i = \kappa_S\left(\dot{\gamma}_i - r_i\left(\sum_{j=1}^{n}\left(\mathbf{G}^{-1}\right)_{ij}\dot{\gamma}_j - \left(\mathbf{G}^{-1}\mathbf{c}\right)_i\right), r_i\right) \qquad i \in P_{T3}. \tag{46}$$

In summary, the contact behaviour of a non-smooth mechanical system can be described by a set of n non-linear equations. These equations can either be formulated in terms of the forces $\boldsymbol{\lambda}$ or in terms of the accelerations $\dot{\boldsymbol{\gamma}}$. In analogy to the inclusions we call them $\boldsymbol{\lambda}$-equations and $\dot{\boldsymbol{\gamma}}$-equations.

A mechanical system with a non-regular matrix $\mathbf{G}$ causes problems. The $\boldsymbol{\lambda}$-equations of such a system might have non-unique solutions, whereas a description of such a system by $\dot{\boldsymbol{\gamma}}$-equations is not even possible. We will therefore only discuss the $\boldsymbol{\lambda}$-equations in the following.

5 Solving the non-linear equations

The $\boldsymbol{\lambda}$-equations (43) and (44) describing the contact behaviour can be solved by different methods. One possibility is to use a Newton–Raphson method [1]. Other possibilities are iterative schemes which are derived from the Jacobi and Gauss–Seidel methods [12], on which we focus in the following.

5.1 Banach fixed point theorem

A system of equations

$$\boldsymbol{\xi} = \mathbf{F}(\boldsymbol{\xi}), \quad \boldsymbol{\xi} \in \mathbb{R}^m, \quad \mathbf{F} : \mathbb{R}^m \to \mathbb{R}^m \tag{47}$$

can be solved by the iterative method

$$\boldsymbol{\xi}^{\nu+1} = \mathbf{F}(\boldsymbol{\xi}^{\nu}), \tag{48}$$

if the function $\mathbf{F}(\boldsymbol{\xi})$ is Lipschitz continuous with a Lipschitz constant L smaller than one (Banach fixed point theorem),

$$||\mathbf{F}(\boldsymbol{\xi}) - \mathbf{F}(\hat{\boldsymbol{\xi}})|| \leq L||\boldsymbol{\xi} - \hat{\boldsymbol{\xi}}||, \quad L < 1, \quad \forall \hat{\boldsymbol{\xi}} \in \mathbb{R}^m. \tag{49}$$

The smaller the Lipschitz constant L, the better the convergence is. Arbitrary norms $||\cdot||$ can be used. We use either the spectral norm

$$||\boldsymbol{\xi}||_2 = \sqrt{\boldsymbol{\xi}^{\top}\boldsymbol{\xi}}, \quad ||\mathbf{A}||_2 = \rho(\mathbf{A}) = \max_h |\mu_h(\mathbf{A})|, \quad ||\mathbf{A}\boldsymbol{\xi}||_2 \leq ||\mathbf{A}||_2||\boldsymbol{\xi}||_2, \tag{50}$$

in which $\mu_h(\mathbf{A})$ are the eigenvalues of $\mathbf{A}$, or the row-sum norm

$$||\boldsymbol{\xi}||_\infty = \max_h |\xi_h|, \quad ||\mathbf{A}||_Z = \max_h \sum_{k=1}^{n} |a_{hk}|, \quad ||\mathbf{A}\boldsymbol{\xi}||_\infty \leq ||\mathbf{A}||_Z||\boldsymbol{\xi}||_\infty. \tag{51}$$

5.2 Jacobi and Gauss–Seidel methods for linear systems

A linear system $\mathbf{A}\mathbf{x} + \mathbf{b} = 0$ with $\mathbf{A} \in \mathbb{R}^{m\times m}$, $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{b} \in \mathbb{R}^m$ can be solved by Jacobi or Gauss–Seidel iterative methods, both based on matrix splitting. The instruction for the Jacobi iteration (J) is

$$x_h^{\nu+1} = -\frac{1}{A_{hh}}\left(\sum_{\substack{k=1\\k\neq h}}^{m} A_{hk}\, x_k^{\nu} + b_h\right). \tag{52}$$

An extension of the Jacobi method is the Jacobi iteration with relaxation (JOR)

$$x_h^{\nu+1} = x_h^{\nu} + \omega_h \left[\underbrace{-\frac{1}{A_{hh}} \left(\sum_{\substack{k=1 \\ k\neq h}}^{m} A_{hk}\, x_k^{\nu} + b_h \right)}_{x_h^{\nu+1}\text{Jacobi}} - x_h^{\nu} \right] = x_h^{\nu} - \frac{\omega_h}{A_{hh}} \left(\sum_{k=1}^{m} A_{hk}\, x_k^{\nu} + b_h \right). \tag{53}$$

If the relaxation parameter ω_h is set equal to one, we arrive at the original Jacobi method without relaxation. Convergence of the JOR method can be treated by the Banach fix point theorem (49). It can only be guaranteed for strictly diagonal dominant matrices $\mathbf{A}$, that is

$$\sum_{\substack{k=1 \\ k\neq h}}^{m} \left| \frac{A_{hk}}{A_{hh}} \right| < 1 \qquad \forall h. \tag{54}$$

Another splitting scheme leads to the Gauss–Seidel iteration method (S),

$$x_h^{\nu+1} = -\frac{1}{A_{hh}} \left(\sum_{k=1}^{h-1} A_{hk}\, x_k^{\nu+1} + \sum_{k=h+1}^{m} A_{hk}\, x_k^{\nu} + b_h \right). \tag{55}$$

The corresponding relaxation method (SOR) reads

$$x_h^{\nu+1} = x_h^{\nu} - \frac{\omega_h}{A_{hh}} \left(\sum_{k=1}^{h-1} A_{hk}\, x_k^{\nu+1} + \sum_{k=h}^{m} A_{hk}\, x_k^{\nu} + b_h \right). \tag{56}$$

Convergence of the SOR method can only be guaranteed for positive definite matrices $\mathbf{A}$.

5.3 Jprox method

We present now a Jacobi like iterative instruction to solve the non-linear system of $\boldsymbol{\lambda}$-equations, which we call Jprox method. For the $\boldsymbol{\lambda}$-equation (34) of a unilateral constraint, we propose the scheme

$$\lambda_i^{\nu+1} = \operatorname{prox}_{\mathbb{R}_0^+} \left(-\frac{1}{G_{ii}} \left(\sum_{\substack{j=1 \\ j\neq i}}^{n} \mathbf{G}_{ij}\, \boldsymbol{\lambda}_j^{\nu} + c_i \right) \right). \tag{57}$$

Note that this instruction consists of the Jacobi method for a linear system $\mathbf{G}\,\boldsymbol{\lambda} + \mathbf{c} = 0$ combined with a projection onto the set $\mathbb{R}_0^+$. This means that we treat the unilateral constraint like a bilateral constraint and add a projection. The instruction (57) is identical to the iterative scheme used in [3] to solve linear complementarity problems

$$\dot{\gamma} = \mathbf{G}\boldsymbol{\lambda} + \mathbf{c}, \quad \dot{\gamma} \geq 0, \quad \boldsymbol{\lambda} \geq 0, \quad \dot{\gamma}^{\top}\boldsymbol{\lambda} = 0. \tag{58}$$

In a similar way, we choose for the spatial friction constraint (39) the scheme

$$\boldsymbol{\lambda}_i^{\nu+1} = \operatorname{prox}_S \left(-\frac{1}{\alpha_i} \left(\sum_{\substack{j=1 \\ j\neq i}}^{n} \mathbf{G}_{ij}\, \boldsymbol{\lambda}_j^{\nu} + \mathbf{c}_i - \mathbf{B}_{ii}\boldsymbol{\lambda}_i^{\nu} \right) \right). \tag{59}$$

This instruction is quite similar to the Jacobi method (52). Note however, that both components of the friction force $\boldsymbol{\lambda}_i$ are simultaneously iterated in (59). In terms of the notation used in (52), this iteration would simultaneously affect two components x_h and x_{h+1}. Moreover, the factor α_i in (59) can not be regarded as

two successive diagonal entries A_{nn}, A_{n+1n+1}, but is related to the problem through (37). As a consequence, also the right-hand side of (59) contains $\boldsymbol{\lambda}_i$. This is different from the original Jacobi method (52), in which x_h occurs only on the left-hand side.

5.4 JORprox method

We use (43), (44) to derive an iterative instruction for the $\boldsymbol{\lambda}$-equations, which corresponds to a Jacobi relaxation method,

$$\lambda_i^{\nu+1} = \text{prox}_{\mathbb{R}_0^+}\left(\lambda_i^{\nu} - r_i\left(\sum_{j=1}^{n} \mathbf{G}_{ij}\,\boldsymbol{\lambda}_j^{\nu} + c_i\right)\right) =: \text{prox}_{\mathbb{R}_0^+}\left(F_i\left(\boldsymbol{\lambda}^{\nu}\right)\right) \quad i \in P_N, \tag{60}$$

$$\boldsymbol{\lambda}_i^{\nu+1} = \text{prox}_{S}\left(\boldsymbol{\lambda}_i^{\nu} - r_i\left(\sum_{j=1}^{n} \mathbf{G}_{ij}\,\boldsymbol{\lambda}_j^{\nu} + \mathbf{c}_i\right)\right) =: \text{prox}_{S}\left(\mathbf{F}_i\left(\boldsymbol{\lambda}^{\nu}\right)\right) \quad i \in P_{T3}. \tag{61}$$

Both instructions are composed of the JOR method combined with an additional projection, and are therefore addressed as the JORprox method. Note that the factor r_i in (60) corresponds to the relaxation factor ω_h in (53), that is $r_i = \omega_h / G_{hh}$, where h is the row in $\mathbf{G}$ which belongs to the unilateral constraint i. If i is an spatial friction constraint, it affects two rows h and $h+1$ in $\mathbf{G}$. Thus the JORprox (61) merges two JOR method instructions (53) with relaxation factors ω_h and ω_{h+1}. The two relaxation factors cannot be chosen independently, because of

$$r_i = \frac{\omega_h}{G_{hh}} = \frac{\omega_{h+1}}{G_{(h+1)(h+1)}}. \tag{62}$$

Methods similar to JORprox are used by [1,5,9].

5.5 Convergence of the JORprox method

In this section we show that the JORprox method converges if the underlying JOR method does. The proof uses the fact that the prox_C mapping is a contraction. In particular, it holds for $\text{prox}_{\mathbb{R}_0^+}$ and $F_i(\boldsymbol{\lambda})$ as defined in (60) that

$$||\text{prox}_{\mathbb{R}_0^+}(F_i(\boldsymbol{\lambda})) - \text{prox}_{\mathbb{R}_0^+}(F_i(\hat{\boldsymbol{\lambda}}))||_2 \leq ||F_i(\boldsymbol{\lambda}) - F_i(\hat{\boldsymbol{\lambda}})||_2 \tag{63}$$

because the derivative of $\text{prox}_{\mathbb{R}_0^+}$ is always equal to or less than one, see Fig. 3. The mapping prox_S together with $\mathbf{F}_i(\boldsymbol{\lambda})$ from (61) satisfies

$$||\text{prox}_S(\mathbf{F}_i(\boldsymbol{\lambda})) - \text{prox}_S(\mathbf{F}_i(\hat{\boldsymbol{\lambda}}))||_2 \leq ||\mathbf{F}_i(\boldsymbol{\lambda}) - \mathbf{F}_i(\hat{\boldsymbol{\lambda}})||_2 \tag{64}$$

which is visualized in Fig. 4. Note that the vector from $\mathbf{F}(\hat{\boldsymbol{\lambda}})$ to $\mathbf{F}(\boldsymbol{\lambda})$ reduces in length after the projection of $\mathbf{F}(\hat{\boldsymbol{\lambda}})$ and $\mathbf{F}(\boldsymbol{\lambda})$ to the set S has been performed. Note also that (64) applies for a contact with associated friction, i.e. friction that does not depend on the contact's normal load. In this case, the set S of admissible tangential forces is constant as it has been presumed in (64).

Inequalities of type (63) or (64) may be combined to state global estimations, which are necessary for the multi-contact case. If, for example, one unilateral contact (63) and one spatial friction constraint (64) is considered, one obtains because of the contractivity of prox

$$\left\| \begin{pmatrix} \text{prox}_{\mathbb{R}_0^+}(F_1(\boldsymbol{\lambda})) - \text{prox}_{\mathbb{R}_0^+}(F_1(\hat{\boldsymbol{\lambda}})) \\ \text{prox}_S(\mathbf{F}_2(\boldsymbol{\lambda})) - \text{prox}_S(\mathbf{F}_2(\hat{\boldsymbol{\lambda}})) \end{pmatrix} \right\|_2 \leq \left\| \begin{pmatrix} F_1(\boldsymbol{\lambda}) - F_1(\hat{\boldsymbol{\lambda}}) \\ \mathbf{F}_2(\boldsymbol{\lambda}) - \mathbf{F}_2(\hat{\boldsymbol{\lambda}}) \end{pmatrix} \right\|_2 := \|\mathbf{F}(\boldsymbol{\lambda}) - \mathbf{F}(\hat{\boldsymbol{\lambda}})\|. \tag{65}$$

Furthermore, $\mathbf{F}(\boldsymbol{\lambda})$ can be written by (60), (61) in the form $\mathbf{F}(\boldsymbol{\lambda}) = \boldsymbol{\lambda} - \mathbf{R}(\mathbf{G}\boldsymbol{\lambda} + \mathbf{c})$, in which $\mathbf{R}$ is a diagonal matrix with entries r_i,

$$\mathbf{R} = \begin{pmatrix} r_1 & 0 & 0 \\ 0 & r_2 & 0 \\ 0 & 0 & r_2 \end{pmatrix} = \begin{pmatrix} \frac{\omega_1}{G_{11}} & 0 & 0 \\ 0 & \frac{\omega_2}{G_{22}} & 0 \\ 0 & 0 & \frac{\omega_3}{G_{33}} \end{pmatrix}. \tag{66}$$

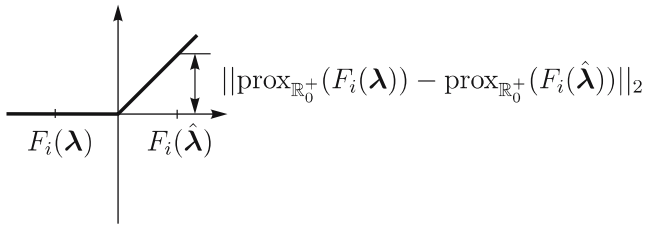


Fig. 3 The function $\text{prox}_{\mathbb{R}_0^+}$

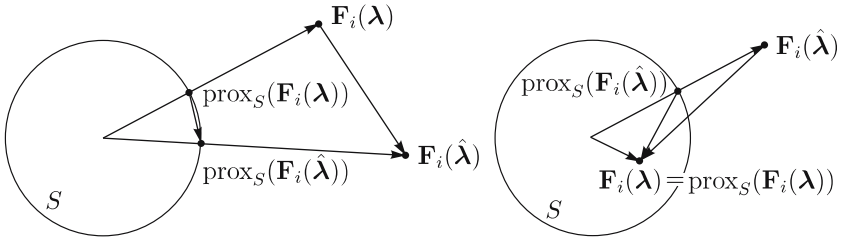


Fig. 4 Graphical interpretation of inequality (64)

We therefore obtain from (50) the estimation

$$\|\mathbf{F}(\boldsymbol{\lambda}) - \mathbf{F}(\hat{\boldsymbol{\lambda}})\|_2 \leq \underbrace{||\mathbf{I} - \mathbf{R}\,\mathbf{G}||_2}_{L}\ ||\boldsymbol{\lambda} - \hat{\boldsymbol{\lambda}}||_2 \tag{67}$$

which proves together with (65) that the JORprox method satisfies the Banach fixed point theorem if the JOR method does. To guarantee the convergence of the underlying JOR method, the Lipschitz constant L in (67) must be less than one. This can be expressed by a condition on the values of ω_h, which we derive by estimating the spectral norm with the help of the max-row norm,

$$L = ||\mathbf{I} - \mathbf{R}\,\mathbf{G}||_2 \leq \underbrace{||\mathbf{I} - \mathbf{R}\,\mathbf{G}||_\infty}_{\tilde{L}} = \max_h \left(|1 - \omega_h| + \omega_h \sum_{\substack{k=1 \\ k \neq h}}^{m} \left| \frac{G_{hk}}{G_{hh}} \right| \right) < 1. \tag{68}$$

In Fig. 5, the estimation $\tilde{L}$ of the Lipschitz constant L is shown as a function of ω_h, together with the two summands that build it up. The left diagram corresponds to a matrix $\mathbf{G}$ which is strictly diagonal dominant. In this case, a minimum Lipschitz constant $\tilde{L}$ is reached for $\omega_h = 1$. In the right diagram, the matrix $\mathbf{G}$ is not strictly diagonal dominant. As a consequence, the estimate $\tilde{L}$ of the Lipschitz constant continuously increases from one. Although convergence cannot be guaranteed for this case, the numerical scheme might still work. This results from the fact that $\tilde{L}(\omega_h)$ is close to one for small values of ω_h. However, choosing ω_h small does *not* solve the problem, because this choice influences the stop criterion of the algorithm, which is based on a relative and absolute error estimate: The choice $\omega_h = 0$ causes $r_i = 0$ by (66). As a consequence, the iterative instructions (60), (61) for inclusion i become independent from all other inclusions j, and the algorithm terminates already after the first loop, because after it the values of λ_i are not changed anymore. Note that the parameter r_i controls the influence of the j-th inclusion on the i-th inclusion.

Based on these observations, we suggest the following choice of parameter r_i: if the matrix $\mathbf{G}$ is strictly diagonal dominant, we set

$$r_i = \frac{1}{G_{hh}}, \tag{69}$$

in which h denotes the row in $\mathbf{G}$ which belongs to constraint i. In the case of spatial friction, h addresses any of the associated rows in $\mathbf{G}$, preferably the one which minimizes r_i. If $\mathbf{G}$ is not strictly diagonal dominant, then r_i has to be chosen small. A good empiric criterion is

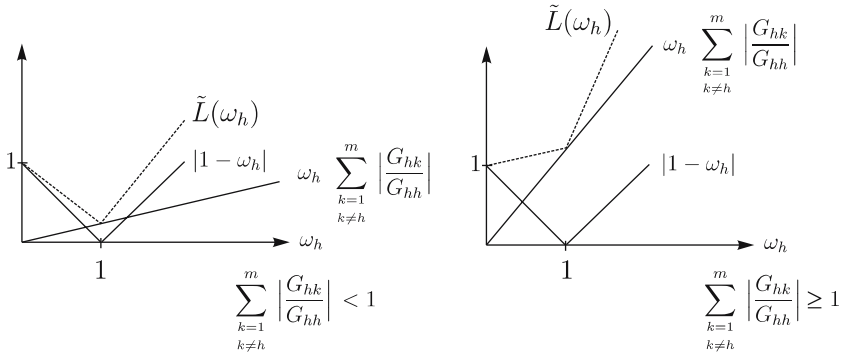


Fig. 5 Minimizing the Lipschitz constant $\tilde{L}(\omega_h)$ (68). In the *left figure*, $\mathbf{G}$ is strictly diagonal dominant, and a minimum can be reached for $\omega_h = 1$. In the *right figure*, $\mathbf{G}$ is not strictly diagonal dominant, and the Lipschitz constant $\tilde{L}(\omega_h)$ is always greater than one

$$\frac{1}{r_i} = \sum_{k=1}^{m} |G_{hk}| \,. \tag{70}$$

If the diagonal elements predominate, then the criteria (69) and (70) become similar.

5.6 Interpretation

The Jprox method can be interpreted in the following way (Fig. 6): in order to calculate the constraint force $\boldsymbol{\lambda}_i^{\nu+1}$ of the constraint i, we assume that all other constraint forces $\boldsymbol{\lambda}_j^{\nu}$ are already known. If the constraint i is unilateral or frictional with $\mathbf{B}_{ii} = 0$, it can be solved straightforward. If the constraint i is a spatial friction constraint with $\mathbf{B}_{ii} \neq 0$, then a better approximation is gained. Then the next constraint $i+1$ is treated under the same assumption that all other constraint forces are known. The calculation of the constraint force $\boldsymbol{\lambda}_{i+1}^{\nu+1}$ uses $\boldsymbol{\lambda}_i^{\nu}$, thus it does not utilize the fact that $\boldsymbol{\lambda}_i^{\nu+1}$ has already been computed. The update of the forces $\boldsymbol{\lambda}^{\nu} \rightarrow \boldsymbol{\lambda}^{\nu+1}$ is not performed until all constraints have been treated. The parameter r_i in the JORprox is equivalent to a relaxation factor and determines the influence of the constraints j on the constraint i.

5.7 SORprox method

It is also possible to solve the $\boldsymbol{\lambda}$-equations with a SOR like method. In contrast to (60), (61), the iterative procedure is now

$$\lambda_i^{\nu+1} = \operatorname{prox}_{\mathbb{R}_0^+}\left(\lambda_i^{\nu} - r_i\left(\sum_{j=1}^{j<i} \mathbf{G}_{ij}\boldsymbol{\lambda}_j^{\nu+1} + \sum_{j=i}^{n} \mathbf{G}_{ij}\boldsymbol{\lambda}_j^{\nu} + c_i\right)\right) \quad i \in P_N, \tag{71}$$

$$\boldsymbol{\lambda}_i^{\nu+1} = \operatorname{prox}_{S}\left(\boldsymbol{\lambda}_i^{\nu} - r_i\left(\sum_{j=1}^{j<i} \mathbf{G}_{ij}\boldsymbol{\lambda}_j^{\nu+1} + \sum_{j=i}^{n} \mathbf{G}_{ij}\boldsymbol{\lambda}_j^{\nu} + \mathbf{c}_i\right)\right) \quad i \in P_{T3}. \tag{72}$$

and deviates from (60), (61) in the immediate update of the calculated forces within the loop. We call the instructions (71) and (72) the SORprox method. It is presumable that the convergence of these instructions is the same as in the SOR method, because the prox_C mapping is a contraction. Note that the SOR method converges for positive definite matrices $\mathbf{G}$. This is the case if all possible constraint states do not cause an underdetermined system. If the constraints can act on the system in a way that it becomes underdetermined, then $\mathbf{G}$ is only positive semidefinite and convergence cannot be guaranteed. The SORprox method without relaxation (Sprox) can be interpreted in the following way (Fig. 6).

All constraints are solved by the assumption that all other contact forces are known. By contrast to the Jprox method, the constraint forces are updated after each calculation, thus the calculation of $\boldsymbol{\lambda}_{i+1}^{\nu+1}$ uses $\boldsymbol{\lambda}_i^{\nu+1}$ instead of $\boldsymbol{\lambda}_i^{\nu}$. This corresponds to a successive solution of the individual constraints [11].

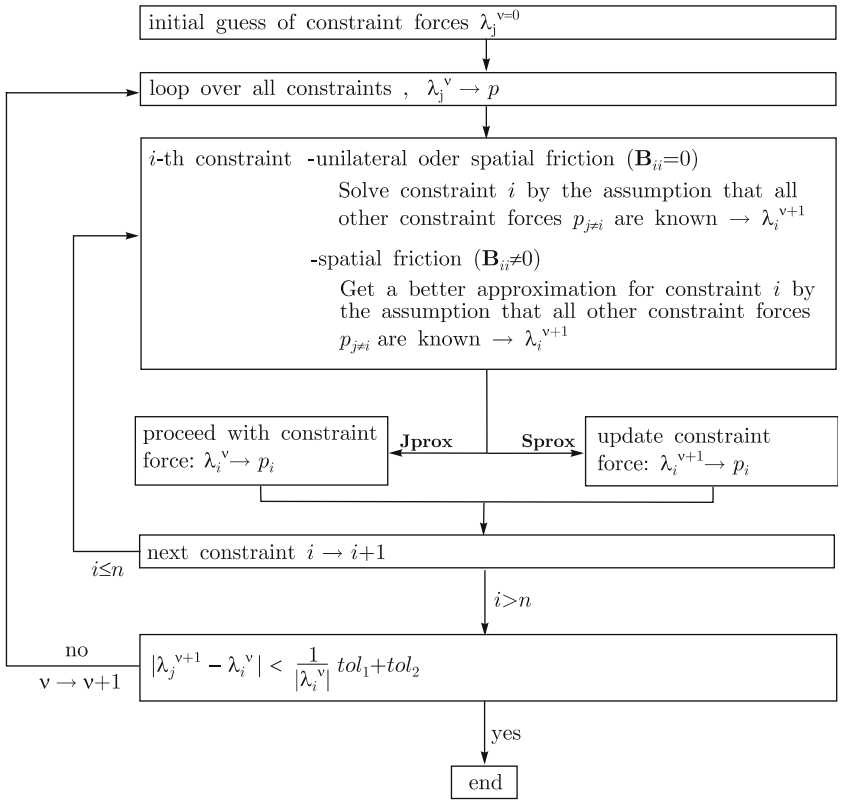


Fig. 6 Interpretation of the Jprox and Sprox method

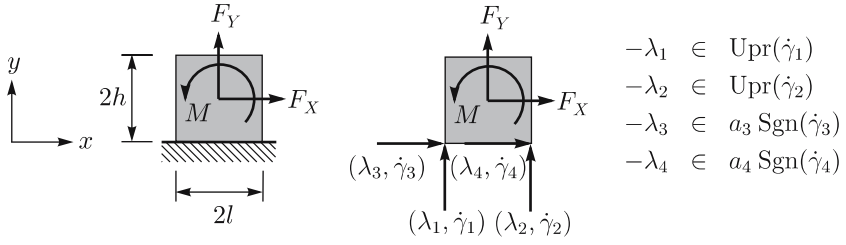


Fig. 7 Planar model of a block on a surface

5.8 Numerical example

In this example we treat a planar system, consisting of a block on a surface with external loads F_X, F_Y, M. The contact between the block and the surface is modeled by two unilateral and two planar frictional constraints, which act on the block's lower corners (Fig. 7). Note that planar friction – as a special case of spatial friction – is simply represented by the filled–in sign function Sgn in (5) and (6), instead of SgnSp. We assume all constraints to be active on velocity level. For such a case, the equations of motion (7) are

$$\underbrace{\begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & J_s \end{pmatrix}}_{\mathbf{M}} \underbrace{\begin{pmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{\varphi} \end{pmatrix}}_{\dot{\mathbf{u}}} - \underbrace{\begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ -\ell & \ell & h & h \end{pmatrix}}_{\mathbf{W}} \underbrace{\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix}}_{\lambda} - \underbrace{\begin{pmatrix} F_x \\ F_y \\ M \end{pmatrix}}_{\mathbf{h}} = 0. \tag{73}$$

We choose as $m = 2.1$, $J_s = 0.018$, $\ell = 0.15$ and $h = 0.05$, for which the $\boldsymbol{\lambda}$-equations (43), (44) become

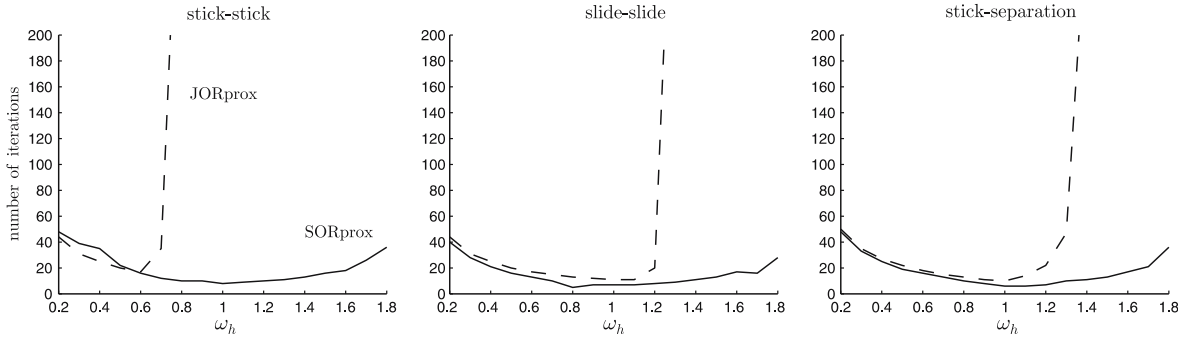


Fig. 8 Convergence of the JORprox (*dashed*) and the SORprox (*solid*) method. The number of iterations is shown as a function of the relaxation parameter ω_h

$$\lambda_1 = \text{prox}_{\mathbb{R}_0^+}(\lambda_1 - r_1\,(1.8\lambda_1 - 0.8\lambda_2 - 0.4\lambda_3 - 0.4\lambda_4 + c_1)), \tag{74}$$

$$\lambda_2 = \text{prox}_{\mathbb{R}_0^+}(\lambda_2 - r_2\,(-0.8\lambda_1 + 1.8\lambda_2 + 0.4\lambda_3 + 0.4\lambda_4 + c_2)), \tag{75}$$

$$\lambda_3 = \text{prox}_{S3}(\lambda_3 - r_3\,(-0.4\lambda_1 + 0.4\lambda_2 + 0.6\lambda_3 + 0.6\lambda_4 + c_3)), \tag{76}$$

$$\lambda_4 = \text{prox}_{S4}(\lambda_4 - r_4\,(-0.4\lambda_1 + 0.4\lambda_2 + 0.6\lambda_3 + 0.6\lambda_4 + c_4)). \tag{77}$$

where c_i contains the terms $\mathbf{c} = \mathbf{W}^\top\mathbf{M}^{-1}\mathbf{h}$. The matrix $\mathbf{G} = \mathbf{W}^\top\mathbf{M}^{-1}\mathbf{W}$ is neither strictly diagonal dominant nor positive definite, but only positive semidefinite. Thus, we cannot guarantee convergence, and the system (74), (75), (76), (77) might have non-unique solutions for the forces $\boldsymbol{\lambda}$.

As an example, consider the case that the block remains in the stick state, which can be realized by moderate or vanishing loading (F_X, F_Y, M). For this situation one can show that non of the prox_C-functions project on their sets ($\text{prox}_C(x) \equiv x$). As a consequence, (74), (75), (76), (77) reduce to a set of *linear* equations which are not independent due to the semi-definiteness of $\mathbf{G}$, and the contact forces can not be determined uniquely. However, we are already satisfied if the algorithm returns just one of those solutions, because in this example all possible force distribution lead to the same (unique) generalized accelerations $\dot{\mathbf{u}}$. If one of the contacts is going to slide or open, then the contact forces are expected to be unique. In these situations, the solution $\boldsymbol{\lambda}$ of the contact problem (74), (75), (76), (77) is at a point at which some of the prox_C-functions really project, which makes the *non*-linear equations of the system (74), (75), (76), (77) independent.

The non-linear equations (74), (75), (76), (77) are solved by the JORprox and the SORprox method. In Fig. 8 the number of iteration steps is shown in function of the relaxation factor ω_h. It is assumed that all constraints have the same relaxation parameter. Convergence cannot be guaranteed, but the iteration converges anyway for some ω_h. In the first plot both unilateral constraints are closed and both friction constraints stick. The solution is non-unique. In the plot we see the number of iteration steps of the JORprox method (dashed) and of the SORprox method (solid). In the second plot both unilateral constraints are closed and both friction constraints slide. Further we have in the third plot one closed unilateral constraint. The corresponding friction constraint sticks. The other unilateral constraint opens.

6 Augmented Lagrangian

It has been shown in Sect. 4 that the inclusions describing unilateral and frictional contact behaviour can be written as equations. In the following we will briefly discuss how those equations are connected to optimization theory. As described in [2,10], a constrained optimization problem can be turned into an unconstrained saddle point problem of the so-called augmented Lagrangian function. We review this mathematical theory and apply it afterwards to mechanics by using the principle of least constraints and it's dual principle. We show that the non-linear equations (43), (44) and (46) describing the contact behaviour are the necessary and sufficient conditions for the saddle point of the augmented Lagrangian function when applied to mechanics.

6.1 Equality constraints

We consider a strictly convex cost function $f(\mathbf{x})$ together with an affine equality constraint $g(\mathbf{x}) = 0$. Associated with the optimization problem

$$\min_x \; f(\mathbf{x}) \quad \text{for} \quad g(\mathbf{x}) = 0 \tag{78}$$

is the augmented Lagrangian function [2]

$$L_a(\mathbf{x}, \mu) = \; f(\mathbf{x}) + \mu \, g(\mathbf{x}) + \frac{r}{2} \, g^2(\mathbf{x}). \tag{79}$$

The factor μ is the Lagrange multiplier known from the ordinary Lagrangian, and r an additional penalty parameter. Finding the saddle point of the augmented Lagrangian function, i.e.

$$\min_x \max_\mu \; L_a(\mathbf{x}, \mu), \tag{80}$$

is equivalent to solving the optimization problem (78).

6.2 Inequality constraint

An optimization problem with an inequality constraint

$$\min_x \; f(\mathbf{x}) \quad \text{for} \quad g(\mathbf{x}) \leq 0 \tag{81}$$

is reformulated with the help of a slack variable ν to obtain

$$\min_x \; f(\mathbf{x}) \quad \text{for} \quad g(\mathbf{x}) + \nu = 0 \,, \quad \nu \geq 0. \tag{82}$$

The associated augmented Lagrangian function is [2]

$$L_a(\mathbf{x}, \mu, \nu) = \; f(\mathbf{x}) + \underbrace{\mu \, (g(\mathbf{x}) + \nu) + \frac{r}{2} \, (g(\mathbf{x}) + \nu)^2}_{P(g(\mathbf{x}), \nu, \mu)}, \tag{83}$$

and the corresponding saddle point problem

$$\min_{x, \nu \geq 0} \max_\mu \; L_a(\mathbf{x}, \mu, \nu). \tag{84}$$

The term $P(g(\mathbf{x}), \nu, \mu)$ in (83) does not depend on $f(\mathbf{x})$ and can be minimized separately with respect to $\nu \geq 0$. This leads to the unconstrained saddle point problem [2]

$$\min_x \max_\mu \; \hat{L}_a(\mathbf{x}, \mu) = \min_x \max_\mu \left[f(\mathbf{x}) - \frac{\mu^2}{2r} + \frac{1}{2r} \operatorname{prox}^2_{\mathbb{R}^+_0}(\mu + r \, g(\mathbf{x})) \right]. \tag{85}$$

Note that $\hat{L}_a(\mathbf{x}, \mu)$ is continuously differentiable, because the continuous function $\operatorname{prox}_{\mathbb{R}^+_0}$ appears quadratic. The necessary conditions for the saddle point are

$$\frac{\partial \hat{L}_a(\mathbf{x}, \mu)}{\partial \mathbf{x}} = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} + \frac{\partial g(\mathbf{x})}{\partial \mathbf{x}} \operatorname{prox}_{\mathbb{R}^+_0}(\mu + r \, g(\mathbf{x})) = 0, \tag{86}$$

$$\frac{\partial \hat{L}_a(\mathbf{x}, \mu)}{\partial \mu} = -\frac{\mu}{r} + \frac{1}{r} \operatorname{prox}_{\mathbb{R}^+_0}(\mu + r \, g(\mathbf{x})) = 0. \tag{87}$$

Equations (85), (86), (87) hold in this form not only for inequality constraints on $\mathbf{x}$ as in (81), but also for any restrictions on the Lagrangian multiplier μ that can be expressed together with a linear-affine function $g(\mathbf{x})$ as a $\operatorname{prox}_C(\cdot)$-condition. This includes in particular the two- and three-dimensional friction law in mechanics, for which the sets C has to be chosen as the friction interval and the friction disk, respectively.

6.3 Mechanical systems with unilateral constraints

The framework of the augmented Lagrangian theory can be used in dynamics in connection with the principle of least constraints. A mechanical system with one unilateral constraint is represented on acceleration level by the optimization problem

$$\min_{\dot{\mathbf{u}}} Z(\dot{\mathbf{u}}) = \frac{1}{2}(\dot{\mathbf{u}} - \dot{\mathbf{u}}_u)^\top \mathbf{M}\,(\dot{\mathbf{u}} - \dot{\mathbf{u}}_u) \quad \text{for} \quad -\dot{\gamma}_i = -\mathbf{w}_i^\top \dot{\mathbf{u}} - \zeta_i \le 0. \tag{88}$$

The cost function Z is called the "Zwang"[1] of the system. The accelerations of the constrained and unconstrained system are denoted by $\dot{\mathbf{u}}$ and $\dot{\mathbf{u}}_u$, respectively, where $\dot{\mathbf{u}}_u = \mathbf{M}^{-1}\mathbf{h}$. By applying conditions (86) and (87) we obtain

$$\mathbf{M}\dot{\mathbf{u}} - \mathbf{h} - \mathbf{w}_i \,\mathrm{prox}_{\mathbb{R}_0^+}(\lambda_i - r\dot{\gamma}_i) = 0, \tag{89}$$

$$-\lambda_i + \mathrm{prox}_{\mathbb{R}_0^+}(\lambda_i - r\dot{\gamma}_i) = 0, \tag{90}$$

where the Lagrange multiplier μ corresponds to the constraint force λ_i. Elimination of $\mathrm{prox}_{\mathbb{R}_0^+}$ from (89) yields

$$\mathbf{M}\dot{\mathbf{u}} - \mathbf{h} - \mathbf{w}_i\,\lambda_i = 0, \tag{91}$$

$$-\lambda_i + \mathrm{prox}_{\mathbb{R}_0^+}(\lambda_i - r\dot{\gamma}_i) = 0. \tag{92}$$

From (91) we recognize the equations of motion for a system with one constraint. Equation (92) expresses the set-valued force law

$$-\lambda_i \in \mathrm{Upr}(\dot{\gamma}_i) \quad \Leftrightarrow \quad -\lambda_i + \mathrm{prox}_{\mathbb{R}_0^+}(\lambda_i - r\dot{\gamma}_i) = 0. \tag{93}$$

This important relation follows immediately from (24) when setting $b_i = x_i - r\,y_i$, but can also be gained through convex analysis[2]. To obtain finally the $\boldsymbol{\lambda}$-equation (43) from (91) and (92), one first solves (91) for $\dot{\mathbf{u}}$, then substitute the result in $\dot{\gamma}_i = \mathbf{w}_i^\top \dot{\mathbf{u}} + \zeta_i$, and puts the resulting expression for $\dot{\gamma}_i$ into Eq. (92).

If we start, instead of (88), with the dual principle of least constraints

$$\min_{\lambda_i} \frac{1}{2}\lambda_i^\top \mathbf{G}\lambda_i + c\lambda_i \quad \text{for} \quad -\lambda_i \le 0, \tag{94}$$

in which $\mathbf{G} = \mathbf{w}_i^\top \mathbf{M}^{-1}\mathbf{w}_i$ and $c = \mathbf{w}_i^\top(\mathbf{M}^{-1}\mathbf{h} + \zeta_i)$, and perform all steps from (81), (82), (83), (84), (85), (86), (87), then we arrive finally with the $\dot{\gamma}$- equation (45). In this case, the Lagrangian multiplier μ corresponds to the relative contact acceleration $\dot{\gamma}_i$.

6.4 Box constraints

We consider a constrained optimization problem of the form

$$\min_x f(\mathbf{x}) \quad \text{for} \quad -a \le g(\mathbf{x}) \le a. \tag{95}$$

There are now two possibilities on how to treat this problem: one could either split up the box constraint into two separate inequalities $-g(\mathbf{x}) - a \le 0$ and $g(\mathbf{x}) - a \le 0$ and treat them as in Sect. 6.2, or introduce a slack variable ν such that

$$\min_x f(\mathbf{x}) \quad \text{for} \quad -a \le g(\mathbf{x}) - \nu \le a\,, \quad \nu = 0. \tag{96}$$

The augmented Lagrangian function for this case is [2]

$$L_a(\mathbf{x}, \mu, \nu) = f(\mathbf{x}) + \underbrace{\mu\,\nu + \frac{r}{2}\nu^2}_{P(\nu,\mu)} \quad \text{for} \quad -a \le g(\mathbf{x}) - \nu \le a. \tag{97}$$

[1] "Zwang": German word for "constraint".

[2] Proof of (93): $-\lambda_i \in \mathrm{Upr}(\dot{\gamma}_i) \Rightarrow -\dot{\gamma}_i \in \mathcal{N}_{\mathbb{R}_0^+}(\lambda_i) \Rightarrow \lambda_i = \mathrm{prox}_{\mathbb{R}_0^+}(\lambda_i - r\dot{\gamma}_i)$.

Minimization of (97) with respect to ν yields the unconstrained saddle point problem [2]

$$\min_x \max_\mu \hat{L}_a(\mathbf{x}, \mu) = \min_x \max_\mu \left[f(\mathbf{x}) - \frac{\mu^2}{2r} + \frac{1}{2r} \kappa_S^2(\mu + rg(\mathbf{x}), r) \right], \tag{98}$$

in which the the function κ_S takes into account the two-dimensional special case of (23). Since $\hat{L}_a(\mathbf{x}, \mu)$ is continuously differentiable, we obtain for the saddle point

$$\frac{\partial \hat{L}_a(\mathbf{x}, \mu, r)}{\partial \mathbf{x}} = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} + \frac{\partial g(\mathbf{x})}{\partial \mathbf{x}} \kappa_S(\mu + rg(\mathbf{x}), r) = 0, \tag{99}$$

$$\frac{\partial \hat{L}_a(\mathbf{x}, \mu, r)}{\partial \mu} = -\frac{\mu}{r} + \frac{1}{r} \kappa_S(\mu + rg(\mathbf{x}), r) = 0. \tag{100}$$

6.5 Mechanical systems with planar frictional constraints

The optimization problem (95) can be used to describe planar dry friction within the dual principle of least constraints,

$$\min_{\lambda_i} \frac{1}{2} \lambda_i^\top \mathbf{G} \lambda_i + c\lambda_i \quad \text{for} \quad -a \leq -\lambda_i \leq a. \tag{101}$$

After having applied (99) and (100), we arrive at

$$\mathbf{G}\lambda_i + c - \kappa_S(\dot{\gamma}_i - r\lambda_i, r) = 0, \tag{102}$$

$$-\dot{\gamma}_i + \kappa_S(\dot{\gamma}_i - r\lambda_i, r) = 0. \tag{103}$$

This is equivalent to

$$\mathbf{G}\lambda_i + c - \dot{\gamma}_i = 0, \tag{104}$$

$$-\dot{\gamma}_i + \kappa_S(\dot{\gamma}_i - r\lambda_i, r) = 0. \tag{105}$$

Elimination of λ_i from (104) and (105) yields the non-linear $\dot{\gamma}$-equation for one planar friction constraint (46). Note that dry friction does not restrict the accelerations $\dot{\mathbf{u}}$, but only the values of the friction force λ_i. As a consequence, the primal principle of least constraints remains *unconstrained* in the presence of dry friction, but has to be modified through addition of support functions that account for the frictional effects [7]. Only in the dual problem (101), the restriction of the friction force remains visible through inequality constraints. Coulomb friction, for which the bounds of the tangential forces depend on the unknown normal load, can *not* be treated by optimization theory. People speak about "quasi optimization" or the like to express that they are solving the tangential sub-problem by methods from optimization theory under the assumption of given normal loads, which in turn are updated within a normal sub-problem under the assumption of given tangential loads.

6.6 Remarks

Optimization problems with several constraints are treated by summing up the functions $P(g(\mathbf{x}), \nu, \mu)$ of the individual constraints in the augmented Lagrangian function. Each constraint has then its own Lagrange multiplier μ_i and penalty factor r_i.

The saddle point of the augmented Lagrangian can also be determined by Uzawa's method [4]. Basically, one first minimizes the augmented Lagrangian function with respect to $\mathbf{x}$ for fixed Lagrange multipliers $\boldsymbol{\mu}$ by using e.g. a steepest gradient method. In a second step, the Lagrange multipliers are updated. For example, the update for a constraint $g(x) \leq 0$ is

$$\mu = \text{prox}_{\mathbb{R}_0^+}(\mu + rg(\mathbf{x})), \tag{106}$$

which corresponds to (87). Roughly spoken, in each minimization step a better solution is gained for $\mathbf{x}$ due to the penalty term. According to this solution $\mathbf{x}$, better Lagrange multipliers $\boldsymbol{\mu}$ can be determined. Attention has to be paid that not all non-smooth constraints can be treated as optimization problems. As already mentioned,

quasi optimization problems are not optimization problems and cannot be solved for example by steepest gradient methods.

Note that instead of combining the two non linear equations (89) and (90) to one non-linear $\boldsymbol{\lambda}$-equation (43), which is solved by an iteration in λ_i, it is also possible to iterate the two nonlinear equations (89) and (90) in λ_i and $\dot{\gamma}_i$:

$$\dot{\gamma}_i^{\nu+1} = \mathbf{G}\,\mathrm{prox}_{\mathbb{R}_0^+}(\lambda_i^\nu - r\dot{\gamma}_i^\nu) + c_i, \tag{107}$$

$$\lambda_i^{\nu+1} = \mathrm{prox}_{\mathbb{R}_0^+}(\lambda_i^\nu - r\dot{\gamma}_i^\nu). \tag{108}$$

An Uzawa-like procedure would be: solve (107) for $\dot{\gamma}_i$ with fixed λ_i, then do one update step for λ_i according to (108).

7 Exact-regularization

Up to now we have shown two different approaches how to obtain a system of equations which describes the behaviour of a non-smooth mechanical system. In Sect. 4 contact inclusions have been turned into equations by analyzing the behaviour of the individual constraints. In Sect. 6, optimization problems related to mechanics have been discussed to arrive at the desired description of the system. We show now a third way how to obtain the same set of equations, but now based on the concept of exact regularization. In Sect. 2 the mechanical system has been described by the equations of motion (7) and set-valued force laws (2), (3), (5) and (6). According to (93), those force laws can directly be transformed into equations and solved together with the equations of motion. The transformation (93) can thereby be regarded as an exact regularization [9], which gives another interpretation of the problem. The main difficulty when dealing with set-valued force law is the fact that a whole set is obtained for some arguments, for example $\mathrm{Upr}(0) = \mathbb{R}_0^-$. One way to avoid this problem is to regularize the set-valued force laws. For example, the set-valued part of the function $\mathrm{Upr}(x)$ can be regularized as a straight line with slope r and intersection $y^\star$ at the ordinate. This leads to a regularized function

$$y = \mathrm{UprRegul}(x) = \mathrm{prox}_{\mathbb{R}_0^+}(y^\star - rx). \tag{109}$$

The function $\mathrm{UprRegul}(x)$ has a unique value y for every function argument x. Of course, the regularized and the set-valued force law are different and do not describe the same behaviour. In Fig. 9, the set-valued force law $-y \in \mathrm{Upr}(x)$ is plotted in light gray, the regularized force law $y = \mathrm{UprRegul}(x)$ law in dashed dark gray.

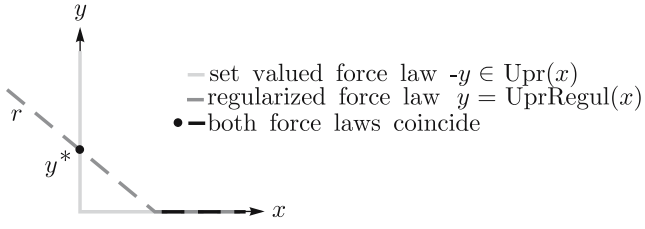


Fig. 9 Exact regularization of the Upr-function

The set-valued force law and the regularized force law coincide if

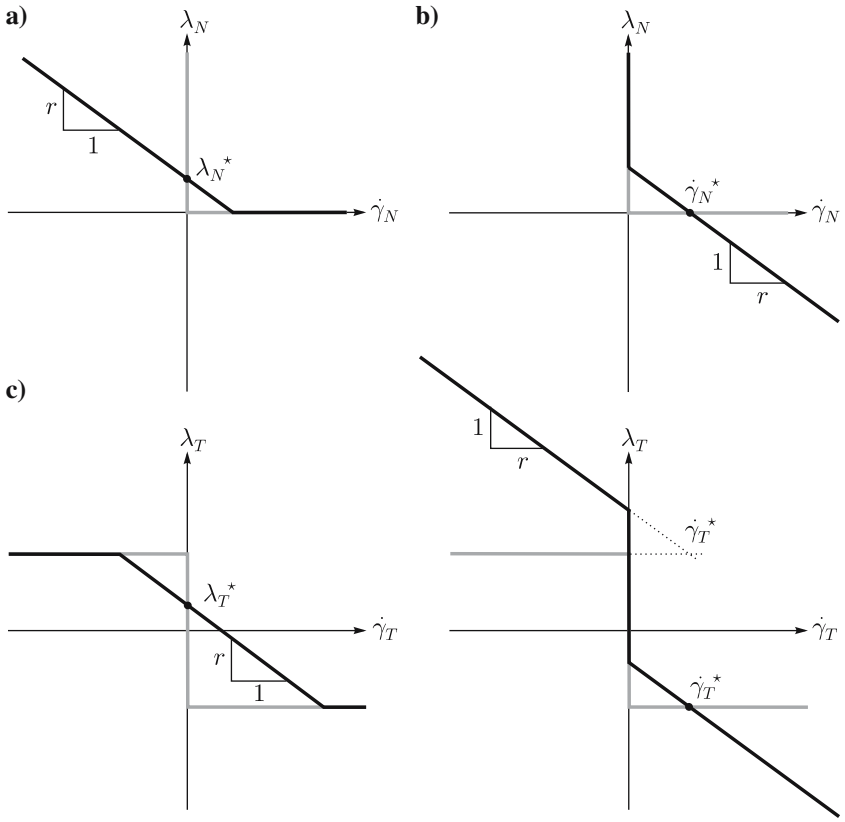
$$(x \geq 0, \quad y = y^\star = 0), \tag{110}$$

$$(x = 0, \quad y = y^\star \geq 0). \tag{111}$$

Thus, every point (x, y) that satisfies

$$y = y^\star = \mathrm{UprRegul}(x) = \mathrm{prox}_{\mathbb{R}_0^+}(y^\star - rx) \tag{112}$$

is a point of the set-valued function $-y \in \mathrm{Upr}(x)$. The idea of exact regularization is to shift the regularized function to the point at which the true solution, i.e. the solution of the set-valued function would be. Figure 10 summarizes the set-valued force laws for unilateral contact and planar friction together with their associated exact regularization.



Inclusion	non-lin. equation	figure
$-\lambda_i \in \mathrm{Upr}(\dot\gamma_i)$	$\lambda_i = \lambda_i^\star = \mathrm{prox}_{\mathbb{R}_0^+}(\lambda_i^\star - r\dot\gamma_i)$	$a)$
$-\dot\gamma_i \in \mathrm{Upr}^{-1}(\lambda_i)$	$\dot\gamma_i = \dot\gamma_i^\star = \mathrm{prox}_{\mathbb{R}_0^+}(\dot\gamma_i^\star - r\lambda_i)$	$b)$
$-\lambda_i \in \mathrm{Sgn}(\dot\gamma_i)$	$\lambda_i = \lambda_i^\star = \mathrm{prox}_S(\lambda_i^\star - r\dot\gamma_i)$	$c)$
$-\dot\gamma_i \in \mathrm{Sgn}^{-1}(\lambda_i)$	$\dot\gamma_i = \dot\gamma_i^\star = \kappa_S(\dot\gamma_i^\star - r\lambda_i, r)$	$d)$

Fig. 10 Non-linear equations describing the set-valued force laws. The regularized force law is shifted to a position such that it intersects the set-valued branch of the contact inclusion in the point at which the solution is

8 Conclusions

In this paper we have discussed three possible methods to obtain the non-linear equations describing the contact behaviour of a non-smooth mechanical system. First, we have merged the equations of motion and the set-valued force laws to obtain inclusions describing the individual non-smooth constraints. By examining the behaviour of one single inclusion, a system of non-linear equations respecting the contact laws have been derived. The non-linear equations have been solved by the JORprox or SORprox iterative method. These solution methods correspond to a successive evaluation of all contacts. The parameter r_i can be interpreted as a relaxation factor and determines the influence of all constraints $j \neq i$ on constraint i. A second possibility to obtain the non-linear equations describing the mechanical system is to set up a constrained optimization problem (principle of least constraints and its dual), which can be turned into an unconstrained saddle point problem of the continuously differentiable augmented Lagrangian function. The conditions for the saddle point lead to the same non-linear equations as above. The parameter r is in this case a penalty factor. A third method is to regularize the set-valued force laws and to shift them into a position at which they coincide with the original ones. This idea is known as exact regularization. The factor r is the slope of the regularized function. Exact regularization links the two main approaches in non-smooth mechanics; the regularized and the set-valued approach.

Appendix: Proof of relations (24), (25), (26)

In the following we prove relations (24), (25), (26). With the help of convex analysis, the functions $\mathrm{Upr}(\cdot)$ and the $\mathrm{SgnSp}^{-1}(\cdot)$ can be expressed as normal cones $\mathcal{N}_{\mathcal{C}}$ to the sets $\mathbb{R}_0^+$ and S [7]. The associated potentials are the indicator functions $\Psi_{\mathbb{R}_0^+}$ and Ψ_S. The function $\mathrm{SgnSp}(\cdot)$ can be expressed as subdifferential of the conjugate potential $\Psi_S^\star(\mathbf{x}_i) = a_i\,||\mathbf{x}_i||_2$. Thus relations (24), (25), (26) can be rewritten as

$$x_i + r\,\mathcal{N}_{\mathbb{R}_0^+}(x_i) \ni b_i \quad \Leftrightarrow \quad x_i = \mathrm{prox}_{\mathbb{R}_0+}(b_i), \tag{113}$$

$$\mathbf{x}_i + r\,\partial\Psi_S^\star(\mathbf{x}_i) \ni \mathbf{b}_i \quad \Leftrightarrow \quad \mathbf{x}_i = \kappa_S(\mathbf{b}_i, r), \tag{114}$$

$$\mathbf{x}_i + r\,\mathcal{N}_S(\mathbf{x}_i) \ni \mathbf{b}_i \quad \Leftrightarrow \quad \mathbf{x}_i = \mathrm{prox}_S(\mathbf{b}_i). \tag{115}$$

Relations (113) and (115) can be treated together by writing in the form

$$\mathbf{x}_i + r\,\mathcal{N}_C(\mathbf{x}_i) \ni \mathbf{b}_i \quad \Leftrightarrow \quad \mathbf{x}_i = \mathrm{prox}_C(\mathbf{b}_i). \tag{116}$$

Reformulating the left part of (116) yields

$$0 \in (\mathbf{x}_i - \mathbf{b}_i) + r\mathcal{N}_C(\mathbf{x}_i), \tag{117}$$

which are for $r > 0$ the optimality conditions of the strictly convex optimization problem

$$\min_{\mathbf{x}_i} \frac{1}{2r}||\mathbf{x}_i - \mathbf{b}_i||_2 + \Psi_C(\mathbf{x}_i). \tag{118}$$

More explicitly, this problem may be stated as

$$\min_{\mathbf{x}_i} \frac{1}{2r}\|\mathbf{x}_i - \mathbf{b}_i\|_2 \quad \text{for} \quad \mathbf{x}_i \in C \tag{119}$$

and determines the nearest point $\mathbf{x}_i$ in C to $\mathbf{b}_i$

$$\mathbf{x}_i = \mathrm{prox}_C(\mathbf{b}_i), \tag{120}$$

which is expressed in the right part of (116).

To prove relation (114), the subdifferential $\partial\Psi_S^\star(\mathbf{x}_i) = a_i\,\partial||\mathbf{x}_i||$ is needed. For $\mathbf{x}_i \neq 0$, $\partial\Psi_S^\star$ is differentiable, and we get

$$a_i\,\partial\|\mathbf{x}_i\| = a_i\,\partial\sqrt{\mathbf{x}_i^\top\mathbf{x}_i} = \frac{a_i\mathbf{x}_i}{\sqrt{\mathbf{x}_i^\top\mathbf{x}_i}} = a_i\,\frac{\mathbf{x}_i}{\|\mathbf{x}_i\|}. \tag{121}$$

For $\mathbf{x}_i = 0$ we use the definition of the subgradient $\mathbf{y}_i \in \partial||0||$, i.e.

$$a_i\|\mathbf{x}_i^\star\| \geq a_i\|0\| + \mathbf{y}_i^\top\,(\mathbf{x}_i^\star - 0) \quad \forall\mathbf{x}_i^\star, \tag{122}$$

and get

$$a_i \geq \mathbf{y}_i^\top\,\frac{\mathbf{x}_i^\star}{||\mathbf{x}_i^\star||} = ||\mathbf{y}_i^\top||\cdot 1\cdot\cos(\phi) \quad \forall\phi\,. \tag{123}$$

Thus, the subdifferential gradient at $\mathbf{x}_i = 0$ consists of all vectors $\mathbf{y}_i$ with length less than or equal to a_i and forms, in this case, precisely the set S,

$$\partial\|0\| \equiv S. \tag{124}$$

With the help of (121) we may now evaluate the left part in (114) for $\mathbf{x}_i \neq 0$, which gives

$$\mathbf{x}_i = \mathbf{b}_i - ra_i\frac{\mathbf{x}_i}{||\mathbf{x}_i||} \quad \Leftrightarrow \quad \mathbf{x}_i\underbrace{(1 + \frac{ra_i}{||\mathbf{x}_i||})}_{>0} = \mathbf{b}_i\,. \tag{125}$$

From the second equation in (125) we see that the two vectors $\mathbf{x}_i$ and $\mathbf{b}_i$ are parallel. It therefore holds that

$$\mathbf{x}_i = \mathbf{b}_i - ra_i \frac{\mathbf{b}_i}{||\mathbf{b}_i||} \quad \Leftrightarrow \quad \mathbf{b}_i \underbrace{\left(1 - \frac{ra_i}{||\mathbf{b}_i||}\right)}_{>0} = \mathbf{x}_i, \tag{126}$$

where the term $(1 - ra_i/||\mathbf{b}_i||)$ has to be greater than zero due to the parallelity of $\mathbf{x}_i$ and $\mathbf{b}_i$. Thus,

$$\mathbf{x}_i = \mathbf{b}_i - ra_i \frac{\mathbf{b}_i}{||\mathbf{b}_i||} \quad \text{if} \quad ||\mathbf{b}_i|| > ra_i \;\Leftrightarrow\; \frac{\mathbf{b}_i}{r} \notin S, \tag{127}$$

which proves the right part in (114) for $\mathbf{x}_i \neq 0$. For $\mathbf{x}_i = 0$ we take into account (124) in the left part of (114), i.e.

$$0 \in \mathbf{b}_i - r\,\partial||0|| = \mathbf{b}_i - r\,S. \tag{128}$$

Thus,

$$\mathbf{x}_i = 0 \quad \text{if} \quad \frac{\mathbf{b}_i}{r} \in S, \tag{129}$$

which proves the right part in (114) for $\mathbf{x}_i = 0$.

Note also that the functions prox_S and the κ_S are linked together by

$$\mathbf{x}_i = \kappa_S(\mathbf{b}_i, r) = \mathbf{b}_i - \text{prox}_{rS}(\mathbf{b}_i). \tag{130}$$

List of Symbols

Upr	unilateral primitives
Sgn	filled-in signum function (one dimensional case, planar friction)
SgnSp	filled-in signum function (two dimensional case, spatial friction)
prox	proximal point function
JORprox	Jacobi relaxation method combined with a projection
SORprox	Gauss-Seidel relaxation method combined with a projection
UprRegul	regularized Upr
$\mathbf{q}$	generalized coordinates
$\mathbf{u}$	generalized velocities
γ_i	relative contact velocity of the i-th constraint (one dimensional, e.g. unilateral)
γ_i	relative contact velocity of the i-th constraint (more dimensional, e.g spatial friction)
γ	vector of all relative contact velocities
λ_i	contact force of the i-th constraint (one dimensional, e.g. unilateral)
$\boldsymbol{\lambda}_i$	contact force of the i-th constraint (more dimensional, e.g. spatial friction)
$\boldsymbol{\lambda}$	vector of all contact forces
$\mathbf{M}$	mass matrix
$\mathbf{h}$	vector of gyroscopic and external forces
$\mathbf{w}_i$	generalized force direction of the i-th constraint (one dimensional, e.g. unilateral)
$\mathbf{W}_i$	generalized force direction of the i-th constraint (more dimensional, e.g. spatial friction)
$\mathbf{W}$	generalized force directions of all constraints
ζ_i	inhomogeneity term of the i-th constraint (one dimensional, e.g. unilateral)
$\boldsymbol{\zeta}_i$	inhomogeneity term of the i-th constraint (more dimensional, e.g spatial friction)
$\boldsymbol{\zeta}$	vector of all inhomogeneity terms
$\mathbf{G}$	$\mathbf{G} = \mathbf{W}^\top \mathbf{M}^{-1} \mathbf{W}$
$\mathbf{c}$	$\mathbf{c} = \mathbf{W}^\top \mathbf{M}^{-1} \mathbf{h}$
$\mathbf{G}_{ij}$	partial matrix of $\mathbf{G}$ linking the i-th and j-th constraint
c_i	partial entry of $\mathbf{c}$ belonging to the i-th constraint (one dimensional, e.g. unilateral)
$\mathbf{c}_i$	partial vector of $\mathbf{c}$ belonging to the i-th constraint (more dimensional, e.g spatial friction)
P_N	set of all unilateral constraints which are active on velocity level
P_{T3}	set of all spatial friction constraints which are active on velocity level

References

1. Alart, P., Curnier, A.: A mixed formulation for frictional contact problems prone to newton like solution methods. Comput Methods Appl Mech Eng. 92(3), 353–375 (1991)
2. Bertsekas, D.P. Constrained Optimization and Lagrange Multiplier Methods. Academic , New York (1982)
3. Cottle, R., Pang, W.: The Linear Complementarity Problem. Academic, New York (1992)
4. Ekeland R.T.: Analyse convexe et Problemes Variationelles. Dunod, Paris (1974)
5. Foerg, M., Ulbrich, H.: Analysis of different time-integration methods applied to a non-smooth industrial problem.In: Proceedings ENOC-2005 Eindhoven (2005)
6. Glocker, C.: Dynamik von Starrkörpersystemen mit Reibung und Stössen, VDI-Fortschrittberichte Mechanik/Bruchmechanik, vol 18/182. VDI-Verlag
7. Glocker, C.: Set-Valued Force Laws–Dynamics of Non-Smooth Systems, Lecture Notes in Applied Mechanics, vol 1. Springer, Berlin Heidelberg New York (2001)
8. Glocker, C., Studer, C.: Formulation and preparation for numerical evaluation of linear complementarity systems in dynamics. Multibody Syst Dyn 13(4),447–463
9. Leine, R.I., Nijmeijer, H.:Dynamics and Bifurcations of Non-smooth Mechanical Systems, Lecture Notes in Applied Mechanics, vol 18. Springer, Berlin Heidelberg New York (2004)
10. Luenberger, D.G.: Linear and Nonlinear Programming. Addison-Wesley, Reading (1984)
11. Moreau, J.J.: Unilateral Contact and Dry Friction in Finite Freedom Dynamics, Non-smooth Mechanics and Applications, CISM Courses and Lectures, vol 302. Springer, Berlin Heidelberg New York (1988)
12. Schwarz, H.R.: Numerische Mathematik. B.G. Teubner (1997)
13. Studer, C., Glocker, C.: Simulation of non-smooth mechanical systems with many unilateral constraints.In: Proceedings ENOC-2005 Eindhoven (2005)